DOCUMENT RESUME

ED 389 377 JC 960 022

AUTHOR Fago, George C.
TITLE Teaching Statistics: Shaping, Fading and Concept Formation.
PUB DATE Mar 95
NOTE 16p.; In: Teaching of Psychology: Ideas and Innovations. Proceedings of the Annual Conference on Undergraduate Teaching of Psychology (9th, Ellenville, NY, March 22-24, 1995); see JC 960 009.
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classroom Techniques; College Freshmen; Computer Uses in Education; Concept Formation; *Cooperative Learning; Course Descriptions; Higher Education; *Introductory Courses; *Learning Laboratories; *Statistical Analysis; *Teaching Methods

ABSTRACT

Based on the assumption that a hands-on approach would enable students to better grasp the usefulness of concepts presented, a course and laboratory component in non-experimental research methodology, statistics, and computer applications for beginning students was implemented at Ursinus College, in Pennsylvania. The course meets 3 hours a week for lecture and 1 hour for the laboratory, and students are primarily freshmen and sophomore psychology majors. The course, and specifically the laboratory experiences, place an emphasis on the principle of "shaping," as students progress in discrete and detailed modules from very elemental instruction in and practice with basic computer commands to analyses of their own psychometric scales using correlation and regression routines. The course also emphasizes the principles of "fading" and "concept formation" in that exact commands for accessing data are provided in lab instructions early in the term, with students merely typing in what they are given, but are gradually eliminated from instructions so that students must use acquired concepts. Lab sessions also offer students an opportunity for collaborative learning, as they cooperate and share insights to help each other. Sample student instructions from laboratory sessions related to frequency and correlation exercises are appended. (BCY)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 389 377

Teaching Statistics: Shaping, Fading and Concept Formation

George C. Fago, Ph.D.

Ursinus College

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Levine

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

IC 960 022

TEACHING STATISTICS: SHAPING, FADING, & CONCEPT FORMATION

George C. Fago, Ph.D.
Ursinus College

For a variety of reasons, both practical and pedagogical, several years ago I designed and implemented a course intended to teach non-experimental research methodology, statistics, and computer applications to introductory students. The course as it has developed assumes little or no background with either computers or statistics. The assumption is that a hands-on approach to these issues in an integrated fashion will enable students to better grasp the utility of these tools and thus prepare them for more formal course work, especially in statistics.

This course is a one semester course which meets for three hours per week for lecture, and one hour per week for lab. Enrollment is limited to a maximum of twenty students, and since it is a required course it is nearly always filled. The students are almost entirely freshmen and sophomores and for the most part are psychology majors. Conceptually the course is divided into three sections: (1) Observational Research, (2) Survey Research, and (3) Correlational Research. In each section, as a collaborative effort, the class designs and carries out an appropriate study. As a part of this process, the issues of sampling (adequacy, size, etc.), operationism, reliability and validity recur repeatedly in class discussion. Students not only are forced to confront these issues at every step in the research design, they also must critique the design in their research reports. This process is

consumes much more time than simple laboratory "demonstration" and so during a term only three studies are actually done.

The laboratory is an integral part of the instructional design. Laboratory meetings are held in a general purpose college computer facility which affords each student an individual monitor, printer, and IBM compatible pc. These machines are 'networked' and students have access to SPSS-Studentware on the network. They are also required to purchase the SPSS-Studentware manual. The students must provide their own diskettes, on which they ware encouraged to save results of their analyses. In order to run the actual lab sessions I employ at least one and preferably two junior/senior students who are veterans of the course and who have demonstrated solid computer skills. As part of their employment, the students make themselves available as computer consultants at other times on an 'as needed' basis.

The first two laboratory meetings are devoted to very elementary materials. Lab one introduces students to computers at what ever level it takes-if need be, how to put a diskette into a drive and how to turn on a machine. The lab assistants are particularly helpful at this point, and I also find that students quickly and spontaneously begin to help each other. I also take time to encourage mutual support. Once everybody is up and running, I walk the students through a number of elementary DOS commands such as DIR, CD, CHKDSK, etc. Students receive handouts which give specific step-by-step instructions which they are to follow and which ask for information to be recorded and/or presents questions to be answered. These are collected when

completed and checked/graded to ensure that each student successfully completes the exercises. During the initial labs, students work more as a group through the exercises. This gives me the opportunity to fill in the background--what is a file, what is a directory, what is DOS--as they actually begin to manipulate and look at these concepts come to life on the monitors. The second lab is also primarily a demonstration--it is a walk through of SPSS. It is frankly intended to be a 'gee whiz', 'bells-and-whistles' experience. Again, by means of handouts and my guidance, we take a introductory tour of SPSS as a data analysis tool. For this lab we use 'canned' data--heights, weights, etc. for the college football team. Students start by simply listing the data file, and then enter a number of spss commands which are simple to simple to execute, intuitively obvious, and yield impressive results. They generate simple frequency tables and bar graphs, and they compute means, etc. At this point the intent is simply to give them some elementary familiarity with the system.

In the following weeks, students actually begin to learn statistics, data analysis, and application software usage. It is in these laboratory experiences that 'shaping,' 'fading,' and 'concept formation' principles are used, and they are utilized both within individual laboratory units as well as across the entire term.

Shaping is inherent in the instructional materials throughout the term. In individual instructional units as well as in the units collectively, the emphasis is always on student production and always moves from very simple to the more complex. For

example, by the fourth lab meeting, results of the first observational study are in and students learn to use SPSS and elementary descriptive statistics in order to describe and summarize their findings. By the end of the term students are analyzing the results of a study which includes a psychometric scale which they developed in addition to other psychometrics, and they are using correlation and regression routines to do it. In a like fashion, each unit itself moves from the simple to the complex in a series of stages. For example, students start by determining the mean of a single variable using the 'frequency' command, then using the 'means' command they determine means for that variable sorting by an independent variable. Over a series of additional steps, they determine means classifying by two variables, and they learn to use addition options with the means command.

It is perhaps trite to describe this as 'shaping'. The 'shaping' is of course in the careful attention to detail. Each instructional unit is very carefully thought through and pre-tested so as to assure virtually complete student success. It is unusual for students to get less than complete credit on the individual laboratory exercises. This serves as a powerful delayed reinforcer in addition to the immediate reinforcement that the system feedback provides. It is very obvious from observation that students find successful completion of these commands to be very reinforcing. At the same time the student is being required in class as well as lab to come to grips with the research--to understand the results and to convey that understanding to others in a research report. The laboratory work is coordinated so that students in the lab actually

work with their own data and are introduced to the procedures which they will use to analyze the data set. In many cases, students leave the lab with significant portions of their data analysis completed.

Fading and concept formation is also implicit within every instructional unit and between instructional units as well. One of the tasks facing the novice data analyst is learning the syntax of the application software and, as anyone who has ever tried teaching software applications knows, this is often a frustrating exercise. In this approach, I try to circumvent this frustration by initially giving such exhaustive and literal instructions that, barring typos, success is assured. Students simply "type" what they are given. Then a series of repetitions with variation follows, which is designed to help the student grasp the 'concept' of the command. At the same time, the student is going through the repetitions, the instructions are becoming progressively more abbreviated. After starting with a literal instruction, students must subsequently carry out varying repetitions with increasingly less literal instructions. As an example, the movement is from "Enter the following command line: MEANS WEIGHT BY YEAR BY POSITION" to "Determine the mean weight for the players broken down by year and position." Fading takes place across labs as well as within them. Initially, for example, students are given the exact commands necessary to access the system and the data file. These commands are 'faded out' so that by the middle of the term all that is needed is to tell the students the name of the data file.

During the course of each lab a good deal of collaborative

learning takes place as well. Students cooperate with each other and readily share insights and offer help to others. I very explicitely encourage this kind of cooperation. Not only does it make my job easier, it also is highly effective. Students are often more sensitive to problems in understanding than I am since they are closer to the problem themselves. They have not 'lost their baseline' as we used to say. In addition lab handouts are returned promptly after checking and students are encouraged to keep them together and to bring them to the lab. They also have access to their texts. Given the emphasis on collaboration in the lab, it is actually easy for a 'student' to coast through the labs passively being the recipient of help from others. In order to prevent this and to encourage active learning students are told at the beginning of the course that they will have a lab final at the end of the term. This final is 'open book and notes;' students confront a novel data file and must carry out a series of commands in order to do a number of analyses on the data. The final, needless to say, is not collaborative. This seems to provide sufficient incentive, since I have yet to have a student get less than a 'B' on the final. And more to the point, I have had a large number of 'computer phobics' who have thanked me for helping them overcome their fears, frustrations, etc., with computers.

Go to overheads.

Field Study Lab Exercise

Directions: In order to prepare for writing your first lab report you are to complete the following demonstrations and procedures using the data file called 'NOSEY.SYS'. This data file is present on the networks in the Olin and Myrin computer labs.

1. Boot the system, and at the system prompt enter 'CD\PSYCY110'. You should then see the system prompt 'G:PSYCH110>'. Then enter 'SPSS-x' using A or B in place of the 'x' depending on the size of your disk.

2. Once you are in SPSS you must declare the NOSEY.SYS file as your active file. When you get the system prompt 'SPSS>' enter the following: GET FILE='NOSEY.SYS'. If you are successful in accessing the file, information regarding the file will be seen at the top of the screen.

3. In order to find out the variable names at the prompt enter DISPLAY. The variable names will be displayed.

You are now ready to begin to analyze, summarize and manipulate the data file. We will begin by demonstrating usages and options available with the frequency command. For each of the steps specified below inspect tables and charts carefully. For some of the frequencies we will request descriptive statistics. In those cases, copy the values for the mean, median, mode, and standard deviation into the space provided.

1. Use the DISPLAY command to list the variable names in the data file. What are these variables?

2. Use the LIST command to list the data.

3. Obtain a frequency table for the variable TOTAL. ENTER FREQUENCIES TOTAL. Be certain you understand what these numbers mean.

4. Obtain a frequency table and barchart for the variable LOOK. ENTER FREQ LOOK /BARCHART. Note here that the command FREQUENCIES has been abbreviated to FREQ.

5. Obtain a frequency table, barchart, and summary statistics for the variable TOTAL. Enter FREQ TOTAL /BARCHART /STATISTICS=ALL.

6. Obtain the summary statistics for the variable LOOK omitting the table and the barchart. Enter FREQ LOOK /FORMAT=NOTABLE /STATISTICS=ALL.

7. Now compute the mean, median, mode, standard deviation, and range for the TOTAL variable and enter the results in the table on the following page.

	MEAN	MEDIAN	MODE	STD. DEV.	RANGE
MIN10					

Notice that in obtaining means using the FREQUENCIES command, we get the average percentage of the variable summed over all conditions. We can also get means broken down by other variables (as you will see in the following steps) by using the means command. This command specifies the variable for which we want the mean and the independent variable to be used for classifying the data. For example, in the means command below a mean will be obtained for variable LOOK broken down by sex; this will produce the means for looking 5 for men and women respectively.

7. Compute the mean number of students looking at the note as a function of sex and enter your results in the table below.

MEANS LOOK BY SEX.

	MEAN
MEN	
WOMEN	

8. Note that there is one problem here. We are using the absolute numbers of men and women. If we find, for example, that more women are or reading the note, is this because women are nosier or is it because there are more women than men in the sample to begin with. What we really need is some relative measure on which to compare men and women. For example, we might want to compare them in terms of the percentages of each sex that are on time or late.We can use SPSS to combine or alter variables according to any algebraicly correct formula. The command to us COMPUTE followed by a new variable name and an algebraic argument or statement. For example, to determine the percentages of students looking at the note out of the total numbers observed we can use the following command:

COMPUTE PERCENT=(LOOK/TOTAL)*100

9. Use the FREQUENCY command to look at the resulting variable.

10. Now use the MEANS command to determine the mean percentages of men and women respectively looking at the note, the total observed, and the percentages looking at the note. Enter your results in the table following after the next paragraph of instructions.

PSYCHOLOGY 110 - METHODS
SURVEY STUDY LAB

1. Using the 'FREQUENCIES' command determine the frequencies and percentages for each type of pattern for the MOST variable. As you already know, this is a standard command which gives results for a single variable.

	FREQUENCY	PERCENT	CUM. PERCENT
Rock			
Rap/hip-hop			
Alternative			
Classical			
Country Western			

2. What do you do if you wish to ask if men and women differ in their music preferences. This question requires that the answers to the MOST variable be tabulated as a function of gender. In order to do this we use the 'CROSSTABS' command. Use this command to look at the relationship of gender to preferred music type (CROSSTABS MOST BY GENDER.). Enter the results in the table below.

PREFERRED MUSIC

	Rock	Rap/hip-hop	Alternative	Classical	Country
MALE					
FEMALE					

A. Which sex appears to favor country western, men or women? _____

3. Other questions could focus on either men or women alone; for example, is a man's preference in music related to what sort of area he grew up in. In this case we would first use the PROCESS IF command to select only men's responses for use in the next analysis. Note that the PROCESS IF command applies only to the very next command. Having made that specification then the analysis command should follow. These commands would appear on individual lines as below when using the 'CROSSTABS' command and the 'OPTIONS' subcommand to determine the relationship between the respondent's musical preference to home area type. The options command gives us additional information: the respective row, column, and total percentages.

```
SELECT IF (SEX=1).
CROSSTABS MOST BY HOME /OPTIONS 3 4 5.
```

Relationship of home area to music preference

	City	Suburb	Rural	Town	Foreign	Total
Rock	R C T					
Rap	R C T					
Alter.	R C T					
Class.	R C T					
Country	R C T					
Total						

A. What percentage of the men prefer Alternative?

_____

B. What percentage of the men are from an urban area?

_____

C. For those respondents who like rap what percentage are from an urban area? _____

4. Sometimes for data analysis the researcher will decide to look at the data in a different way. For example, upon looking at the data we might decide to analyze in terms of preference for rock to preference for all other types of music combined. In order to do this we must use the RECODE command to regroup those variables and we must specify exactly how we wish to recode the variable. An example follows recoding the MOST variable. Enter the command as it is written below:

```
RECODE Most(2 3 4 5=2).
```

From this point on, the MOST variable will consist of two values: 1 or Rock and 2 or Other. If you want to go back to the original coding you would have to reaccess the original data file.

5. Use the 'PROCESS IF', 'CROSSTABS' commands and the 'OPTIONS' and 'STATISTICS' subcommands to determine the relationship between the style women listen to most to the style women prefer to calm themselves down.

```
PROCESS IF SEX=2
CROSSTABS MOST BY CALM /OPTIONS 14 15 /STATISTICS 1.
```

a. In the table below, fill in the observed count, the expected value, and the residuals.

	Pop	Class.	Metal	Alter.	Other
Rock	OB EX RE				
Other	OB EX RE				

b. What is the value of Chi Square? [ ]
What is the probability [ ]

c. Do women who use pop music to calm themselves down also prefer rock music to other forms of music? Explain.

CORRELATION LAB

We will be using a data file called INTIMATE.SYS which contains the data which the class gathered. The data file is well labled so that it is usable as is. In this lab you will be introduced to a number of new analytic techniques which you will also need to use to complete your lab report.

I. LOOKING AT RELATIONSHIPS.

It is often useful to begin a data analysis by looking at the general form of the data. You can accomplish this by using the PLOT command to produce a scatter plot. Once you have accessed the data file using the GET command your first task is to produce a scatter plot of the MASC and FEM scales from the BEM scale. To do this enter the command as shown below.

```
PLOT /PLOT = MASC WITH FEM.
```

Inspect the resulting scatter plot. What does the plot suggest regarding the relationship?

II. COMPUTING RELATIONSHIPS.

As we have already seen, SPSS makes statistical calculations easy. In order to compute the correlation coefficient write the command as below.

```
CORRELATION MASC WITH FEM.
```

Notice that the actual value of the correlation is given as well as its significance level. It is not necessary to write a separate line for each correlation desire. Enter the commands given below. Notice what each command generates and complete the table below.

```
CORRELATION MASC FEM  WITH ESTEEM OPEN.
CORRELATION MASC FEM ESTEEM OPEN.
```

CORRELATIONS

	MASC	FEM	ESTEEM	OPEN
MASC				
FEM				
ESTEEM				
OPEN				

III. ANALYZING SUB-GROUPS.

Sometimes it is of interest, or even necessary, to analyze a single subgroup from the larger sample. For example, computing the mean for the FEM variable for the entire sample gives us an ambigous result since it would be based on both men and women correspondents who might be expected to respond differently to the scale. In order to do this the PROCESS IF command can be used. It

is entered on a single line, and then the desired analysis is requested on the next command line. Enter the commands as below.

```
PROCESS IF (SEX=1).   [ENTER]
COR FEM WITH ESTEEM OPEN.   [ENTER]
```

The PROCESS command applies only to the immediately following analysis. If you wish to do another analysis on that subgroup you must repeat the PROCESS command. Alternately you may use the SELECT IF command. If you use this command the selected subgroup remains your active file. In order to retrieve the full group data you must reissue the GET FILE='FILENAME' command.

IV. SCORING THE BEM SCALE.

As the data file now stands it contains the scores for the masculine, feminine, and neuter subscales of the BEM. What we want to look at is what a persons sex role identification is as measured by the BEM. In order to do this we must combine these scores. Actually, to determine this we use only the MASC AND FEM scales. We determine a median for each, divide correspondents into high and low with respect to these scales, and then do assignment according to the following scheme: LoMASC-LoFEM = UNDIFFERENTIATED, LoMASC-HiFEM = FEMININE, HiMASC-LoFEM = MASCULINE, & HiMASC-HiFEM = ANDROGENOUS. In order to do this enter the commands as below. Be certain that you understand the rationale.

```
COMPUTE SPLMASC=MASC.
COMPUTE SPLFEM=FEM.
FREQ MASC FEM /STAT=MEDIAN.
```

	LOW	HIGH	MEDIAN
MASC			
FEM			

```
RECODE SPLMASC (43 THRU 99=1) (100 THRU 144=2).
RECODE SPLFEM (31 THRU 99=2) (100 THRU 128=4).
```

FEMININE

	LOW - 2	HIGH - 4
LOW - 1	3	5
HIGH - 2	4	6

```
COMPUTE COMBO=SPLMASC + SPLFEM
VALUE LABLES COMBO 3 'UNDIF' 4 'MASC' 5 'FEM' 6 'ANDRO'.
```

Using the MEANS command complete the table of means below.

MEAN ESTEEM SCORE

SEX	COMBO	ESTEEM
FEMALE	UNDIF	
	MASC	
	FEM	
	ANDRO	
MALE	UNDIF	
	MASC	
	FEM	
	ANDRO	